ICONO14

Universidad Rey Juan Carlos

Desafección política y toxicidad en X: una propuesta de clasificación en hashtags electorales

Political Disaffection and Toxicity on X: a Proposal to Classify Electoral Hashtags

Desafeição política e toxicidade em X: uma proposta de classificação em hashtags eleitorais

Ysabel Briceño-Romero[1*]

Liliana Calderón-Benavides[2**]

[1] Universidad Autónoma de Bucaramanga, Colombia
[2] Universidad Autónoma de Bucaramanga, Colombia
[*] Docente e Investigadora del Programa de Comunicación Social, Universidad Autónoma de Bucaramanga, Colombia. Email: ybriceno@unab.edu.co
[**] Directora del Centro de Desarrollo Tecnológico Smart Regions Center, Universidad Autónoma de Bucaramanga, Colombia. Email: mcalderon@unab.edu.co

Recibido: 12/12/2024; **Revisado:** 16/12/2024; **Aceptado:** 07/03/2025; **Publicado:** 19/09/2025

Para citar este artículo: Briceño-Romero, Ysabel; & Calderón-Benavides, Liliana (2025). Desafección política y toxicidad en X: una propuesta de clasificación en hashtags electorales. ICONO 14. *Revista Científica de Comunicación y Tecnologías Emergentes, 23*(1): e2234. https://doi.org/10.7195/ri14.v23i1.2234

Esta obra está bajo una Licencia Creative Commons Atribución 4.0 Internacional

Resumen

Este artículo resume una revisión exploratoria para la detección discursiva de desafección política en Twitter (ahora X) en contextos electorales, partiendo de la clasificación de toxicidad que ofrece la herramienta Perspective. En esta experiencia piloto se aborda, desde un enfoque predominante cuantitativo y descriptivo, la tendencia de mensajes considerados tóxicos dentro de la etiqueta #Elecciones2022 en un universo de 115.493 tuits descargados, cuyos contenidos aluden a contextos relacionados con Costa Rica, Colombia, España, México y Perú. El corpus fue sometido a una puntuación automatizada de toxicidad con la herramienta Perspective y luego se realizó una clasificación manual a los contenidos de una submuestra, según categorías relacionadas con la desafección política. La discusión sugiere que la tipificación discursiva de mensajes con toxicidad severa es eficiente en este caso para la detección de la desafección política, entendiendo que existen evidencias lingüísticas de sentimientos negativos explícitos y, en su mayoría, se reconoce el centro sobre el cual recae el malestar, aunque la diferencia de la muestra en cada contexto país exige réplicas de la metodología para ideas concluyentes. Una revisión longitudinal a largo plazo podría contribuir a la comprensión de la cultura política de contextos específicos, aspecto que puede ser aprovechado en futuras investigaciones para empezar a llenar vacíos con nuevas miradas para la detección de la desafección política en discursos digitales.

Palabras clave

Desafección política; Comunicación política; Hashtag; Elecciones; Democracia; Discurso digital.

Abstract

This article presents an exploratory review for the discursive detection of political disaffection on Twitter (now X) in electoral contexts, based on the toxicity classification provided by the Perspective tool. This pilot study adopts a predominantly quantitative and descriptive approach to analyze the trend of messages classified as toxic within the hashtag #Elecciones2022, using a dataset of 115,493 downloaded tweets

referring to contexts in Costa Rica, Colombia, Spain, Mexico, and Peru. The corpus was subjected to automated toxicity scoring through the Perspective tool, followed by a manual classification of a subsample's content based on categories related to political disaffection. The discussion suggests that the discursive categorization of severely toxic messages is effective in detecting political disaffection in this case, given the presence of explicit linguistic evidence of negative sentiments. However, variations in sample composition across national contexts necessitate methodological replications to draw conclusive insights. A long-term longitudinal review could contribute to a deeper understanding of political culture in specific contexts, an aspect that could be leveraged in future research to introduce new perspectives on detecting political disaffection in digital discourse.

Keywords
Political disaffection; Political communication; Hashtag; Elections; Democracy; Digital discourse.

Resumo

Este artigo resume um percurso conceptual e metodológico para a deteção discursiva do descontentamento político no Twitter (agora X) em contextos eleitorais, a partir da classificação da toxicidade. Nesta experiência piloto, a partir de uma abordagem predominantemente quantitativa, analisa-se a tendência das mensagens consideradas tóxicas dentro do rótulo #Elecciones2022 num universo de 115.493 tweets descarregados, cujos conteúdos se referem a contextos relacionados com a Costa Rica, Colômbia, Espanha, México e Peru. O corpus foi submetido a um score de toxicidade automatizado com a ferramenta Perspective e de seguida foi realizada uma classificação manual do conteúdo de uma amostra, de acordo com o tipo de desafeto. Entendido como um quadro discursivo emocional e negativo em relação ao sistema político, este caso propõe categorias de análise como o descontentamento em relação aos atores políticos, às instituições democráticas, aos traços culturais ou à própria democracia, no meio de conversas em espanhol em X, contextualizadas em processos eleitorais. A discussão sugere o potencial das redes sociais Uma revisão longitudinal de longo prazo poderá contribuir para

a compreensão da cultura política de contextos específicos, aspeto que poderá ser utilizado em futuras pesquisas para começar a preencher lacunas com novas perspetivas sobre abordagens como a espiral do cinismo, com os eixos: construção de notícias digitais, ódio, toxicidade e descontentamento.

Palavras-chave
Insatisfação política; Comunicação política; Hashtag; Eleições; Democracia; Discurso digital.

1. Introducción

La desafección política es entendida como una condición subjetiva de descontento hacia tres posibles dimensiones: el proceso político, los actores políticos y las instituciones democráticas (Torcal y Montero, 2006). Una alta desafección política, provocada por altos índices de corrupción, insatisfacción con resultados de gestión de gobiernos, podría estimular el interés por el cambio, con nuevos movimientos políticos que han transformado el descontento como capital electoral.

Dada las implicaciones que la desafección política, como sentimiento negativo, podría implicar para el apoyo de la democracia, se han desarrollado desde mediados del siglo XX esfuerzos continuos por medirla, con metodologías principalmente centradas en encuestas, entrevistas o grupos focales. Pero con el auge de las redes sociales, algunos estudios han intentado trasladar al entorno digital posibilidades de detección de la expresión de cultura política, desde sentimientos negativos o positivos hacia la democracia, como una forma de interpretar el discurso surgido en redes como Twitter (Briceño-Romero et al., 2022; Arcila-Calderón et al., 2022) dado el potencial estímulo que ocasiona el contexto político electoral, en la publicación de contenidos alusivos a las dimensiones estudiadas por la desafección política. En el marco de la "sociolingüística ciudadana", autores defienden que el lenguaje digital emergente en redes sociales podría ofrecer un vistazo a algunas creencias y actitudes socioculturales históricamente relevantes en el discurso sociopolítico (Bridges, 2021). Estas investigaciones son parte de estos esfuerzos.

En este sentido se propone que la desafección política podría ser detectable en la publicación de mensajes calificados como tóxicos en contextos electorales,

dados los avances de clasificación automatizada que potencian el reconocimiento lingüístico, tomando como centro la insatisfacción en determinados contextos sociopolíticos. Se parte entonces en esta investigación de la idea potencial del discurso para expresar la desafección política, en tanto que se generan en Twitter (X) construcciones lingüísticas identificables como negativas hacia cualquiera de las dimensiones (instituciones democráticas, actores políticos o la democracia misma), dado el peso significativo de las palabras.

A partir del fortalecimiento de experiencias en aprendizaje automático para detectar contenidos negativos en línea se propone entonces explorar la clasificación de mensajes derivados de una etiqueta alusiva a procesos electorales en español en la red social X, a partir de Perspective, una herramienta de código abierto de Jigsaw (incubadora dentro de Google) cuyo objetivo es clasificar el discurso en función de su potencial toxicidad en una conversación.

En este sentido, se han diseñado las siguientes preguntas de investigación:

- ¿Cuál es la tendencia de toxicidad en los contenidos derivados de la etiqueta #Elecciones2022?
- ¿Cuán factible es el reconocimiento de desafección política en los contenidos lingüísticos considerados con toxicidad severa por la herramienta Perspective?

Con este contexto socio técnico, la investigación se propuso una prueba piloto para explorar la detección de desafección política en los mensajes considerados con toxicidad severa en la etiqueta #Elecciones2022, cuyos contenidos en español aludieron en su mayoría a contextos de votaciones en el año 2022 en Colombia, Costa Rica, España y México.

2. Marco referencial

2.1. La desafección política como noción

Un aspecto que ha ocupado parte de la discusión sobre la democracia en el siglo XXI ha sido la expresión de rechazo ciudadano a actores políticos e instituciones, como potencial amenaza a su estabilidad (PNUD, 2022). Para Torcal y Montero (2006) "... los ciudadanos parecen haberse vuelto aún más críticos con respecto

a la forma en que funciona la democracia, el desempeño de las instituciones políticas y las actividades diarias de los actores políticos" (p. 4), lo que despliega una paradójica condición de defensa de la democracia, pero con altos niveles de descontento hacia sus instituciones; esto último, ha sido problematizado bajo la noción de desafección política.

Como parte de una dimensión de la noción de cultura política (Almond y Verba, 1963) o de los grados de satisfacción con la democracia (Valgarðsson y Devine, 2022) la desafección política se relaciona con las expectativas que tiene la ciudadanía con el sistema que le rodea, como un sentimiento negativo traducido en forma general como impotencia, cinismo, falta de confianza (Torcal y Montero, 2006) e, incluso, enojo (Pnud, 2022). Aunque la desafección política ha sido abordada desde diferentes aproximaciones es un consenso identificarla como una condición subjetiva de descontento hacia tres posibles dimensiones: el proceso político, los actores políticos y las instituciones democráticas (Torcal y Montero, 2006). Otra definición, planteada por Megías Collado y Moreno (2022) la resume como un sentimiento negativo hacia los políticos, la política y sus procesos, y hacia un sistema incapaz de hacer frente a las demandas y necesidades de los ciudadanos.

La desafección política alude a un sentimiento subjetivo que tiende a estar estrechamente relacionado con el contexto, como un componente más de la cultura política (Almond y Verba, 1963); por tanto, podría explicarse como una serie de códigos que se instalan lentamente y de forma estructural en el ambiente. Sin embargo, la dinámica rápida de la información podría ayudar a generar marcos coyunturales del contexto, por lo que la tendencia podría variar, como lo defiende Megías Collado (2020).

La metodología para medir la desafección política ha sido discutida y problematizada, no siempre con consensos, principalmente en lo que se refiere a la dificultad de reconocer los diferentes contextos, aunque una diversidad de enfoques ha permitido ampliar la mirada. La escala de la confianza hacia la democracia, por un lado, y en otro extremo la desafección política, ha variado en el tiempo. Fontaneda y Sánchez-Vítores (2018) hacen un balance de lo complejo y variados que han sido los diseños para operacionalizar la desafección política. Toda esta trayectoria en el tema ha llevado a distintas estrategias para abordar la actitud política de la ciudadanía frente a las instituciones y valores de la democracia. La redacción de las preguntas y formas de abordar a la población se han ido matizando

y han surgido subíndices de desapego político y desafección institucional, como expresión operativa de la desafección política (Megías Collado y Moreno, 2022) en una escala negativa de satisfacción hacia la democracia (Valgarðsson y Devine, 2022) o confianza política y sus escalas (Marien, 2011). En la tabla 1 se resumen los esfuerzos por abordar en las encuestas la medición en el tema.

Tabla 1. *Resumen formas de abordar la (des)confianza y (des)afección política, según fuentes encontradas*

Variable	Técnica	Pregunta	Centro / Indicador	Organización / Fuente
(des) confianza política	Encuesta (escala de Lickert)	¿cree usted que puede confiar en que el gobierno de Washington hará lo correcto? ¿El gobierno desperdicia mucho dinero? ¿Está el gobierno dirigido por unos pocos grandes intereses que velan por sí mismos? ¿Son corruptos la mayoría de los políticos? ¿Saben los políticos lo que están haciendo?	Actores políticos	Estudios Electorales de Estados Unidos (ANES) Años sesenta
(des) confianza política	Encuesta (escala numérica)	En una escala del uno al diez ¿cuánto confía usted en?: Parlamento Sistema legal Fuerzas armadas Partidos políticos Candidatos, personas que ejercen el poder político.	Instituciones Actores políticos	Encuesta Social Europea (ESS) Eurobarómetro
(des) confianza política	Encuesta (escala de Lickert)	¿Hasta qué punto respeta usted las instituciones políticas de (país)? ¿Hasta qué punto cree que los derechos básicos de los ciudadanos están bien protegidos por el sistema político de (país)? ¿Hasta qué punto se siente orgulloso de vivir bajo el sistema político de (país)?	Instituciones Cultura país	LAPOP (Noam, Rodríguez, Wilson y Zechmeister, 2023)

(continúa)

Tabla 1. *Resumen formas de abordar la (des)confianza y (des)afección política, según fuentes encontradas (continuación)*

Variable	Técnica	Pregunta	Centro / Indicador	Organización / Fuente
Preferencia sistema político	Encuesta (premisas con respuesta escala de Lickert)	¿Con cuál de las siguientes frases Ud. Está más de acuerdo? "La democracia es preferible a cualquier otra forma de gobierno"; "En algunas circunstancias, un gobierno autoritario puede ser preferible a uno democrático; "A la gente como uno, nos da lo mismo un régimen democrático que uno no democrático".	– Apoyo a democracia – Indiferencia hacia la democracia – Apoyo al autoritarismo	Latinobarómetro (2024)
Desafección política	Encuesta (respuestas con escalas de Lickert)	How interested would you say you are in politics? «How often does politics seem so complicated that you really understand what is going on?	Desapego político	Megías Collado y Moreno (2022)
Desafección política	Encuesta (respuestas con escalas de Lickert)	do you think that politicians in general care what people like you think?	Desafección a actores políticos	Megías Collado y Moreno (2022)
Desafección política	Encuesta (respuestas con escala numérica)	please tell me on a score of 0-10 how much you personally trust each of the institutions I read out. 0 means you do not trust an institution at all, and 10 means you have complete trust	Desafección institucional	Megías Collado y Moreno (2022)

Fuente: *creación propia*.

2.2. Toxicidad y desafección en X: una propuesta

Una de las acciones promovidas por la palabra escrita, desde la comunidad de usuarios en las redes sociales, es la publicación de mensajes tóxicos, un fenómeno que ha estimulado no sólo reflexiones dentro de la comunidad académica, sino también acuerdos y decisiones públicas y organizativas para hacerle frente a esta expresión discursiva en entornos digitales con características distribuidas y no centralizadas. Para finales del siglo XX, una compleja construcción de la opinión pública modeló

premisas que relacionaron la cobertura mediática (generalmente negativa) de los asuntos públicos con esa desconfianza y distancia ciudadana hacia la política y los gobiernos, lo que afianzó hipótesis alrededor de la llamada «espiral del cinismo» (Capella y Jamieson Hall, 1997). No obstante, ante la lógica distribuida de la información y la aparición de las redes sociales, estos enfoques quedan en deuda con el acercamiento a la relación entre marcos noticiosos y discurso de odio en los entornos digitales.

En consenso, el discurso de odio puede entenderse como "una violencia directa o indirecta, verbal o física que consiste en un señalamiento claro hacia una persona o grupo de personas por lo que representan en la sociedad" (Cucurull y Aragó Navarro, 2023). En el entorno digital, esta expresión negativa es reconocida por Qayyum et al. (2023) como mala conducta verbal y toxicidad en redes sociales (OSN, por sus siglas en inglés).

Con el ánimo de hacer frente a las dificultades derivadas del contenido tóxico en redes sociales, el uso de métodos computacionales ha impulsado su detección en plataformas digitales, tanto en la academia como en las grandes organizaciones que dirigen estos entornos.

La exploración del contenido con cargas negativas hacia poblaciones específicas en X, desde procesos automatizados de revisión del discurso, ha estado centrada recientemente en: migración (Arcila-Calderón et al., 2020; Arcila-Calderón et al., 2022); sesgo de género: feminismo (Piñeiro-Otero y Martínez-Rolán, 2021); contexto pandemia (Raghad Alshalan et al., 2020). En el discurso político, se destacan las revisiones centradas en ideología (Amores et al., 2021; Díez-Gutiérrez et al., 2022); y polarización partidista en tiempos de elecciones, como el caso de Estados Unidos (Grimminger y Klinger, 2021) y España (Herrero-Izquierdo et al., 2022). Bollen et al. (2021) llegaron a plantear el análisis a gran escala de seis estados de ánimo en Twitter (tensión, depresión, ira, vigor, fatiga, confusión) para modelar tendencias emotivas predictivas, relacionadas con diferentes contextos sociopolíticos. Horta Ribeiro et al. (2018) experimentan con su estudio una caracterización de "*hateful users*".

Entre las nuevas herramientas de detección del discurso con carga negativa en Twitter, se encuentra Perspective API, relacionada con Google, de código abierto y que emplea aprendizaje automático para identificar lo que se ha determinado como construcciones de texto potencialmente tóxicas.

Perspective es un sistema de clasificación supervisada de texto desarrollado que devuelve puntuaciones de probabilidad de 0 a 1 para los atributos: toxicidad severa, insulto, blasfemia, ataque de identidad, amenaza y sexualidad explícita, en varios idiomas, incluido el español. La categoría toxicidad severa ha sido la que más ha aportado a la definición de discurso de odio y una de las mejores validadas en torno a modelos de etiquetado de contenido abusivo u odioso (Lees et al., 2022). Dados los resultados de entrenamiento de la medida toxicidad, medios de comunicación como *New York Times* y *El País* han estado aplicando esta herramienta para la detección y moderación de comentarios (puede verse con más detalles: https://perspectiveapi.com/case-studies/).

Según se explica en la página oficial de esta herramienta, la toxicidad es uno de los atributos de medición del odio en línea, entendida como "un comentario grosero, irrespetuoso o irrazonable que probablemente te haga abandonar una discusión"[1]. Para la clasificación automatizada del discurso, las pautas de Perspective API sugieren a la investigación en ciencias sociales, entre un puntaje del cero al uno, experimentar con umbrales entre 0.7 y 0.9, la puntuación más alta, para una tipificación de toxicidad severa.

Perspective API ha empezado a ser explorada por la academia. Inicialmente, los esfuerzos se han enfocado en demostrar la vulnerabilidad de la herramienta para reconocer frases potencialmente tóxicas en casos de confusión semántica o modificaciones de palabras clave, en los llamados "ejemplos adversos" (Hosseini et al., 2017; Jain et al., 2018). Esta limitación sigue siendo parte de las dificultades que se generan en la clasificación automatizada de contenidos y su reconocimiento contribuye a desarrollar rutas cada vez más sólidas para mejorar la precisión de clasificadores de comentarios tóxicos como Perspective (Lees et al., 2022). Estudios más recientes, desde un enfoque longitudinal, alcanzan a reconocer la potencialidad de Perspective para identificar en tiempos prolongados perfiles tóxicos que podrían contribuir a mejorar futuras conversaciones en X, como concluyen Qayyun et al. (2023).

Por su parte, esfuerzos también recientes exploran metodologías para interpretar el entorno digital con posibilidades de detección de la expresión de cultura política, desde un acercamiento para clasificar contenidos de apoyo o desafección hacia la democracia en Twitter (Briceño-Romero et al., 2022; Arcila-Calderón et al., 2022). Se sugiere, finalmente, en esta investigación que la desafección política

podría reconocerse como una noción de toxicidad severa tipificada por la herramienta Perspective, una forma potencial de la clasificación del discurso de rechazo a actores y procesos relacionados con el sistema político, en etiquetas relacionadas con entornos electorales, dando continuidad a estudios en estos contextos (Keller y Klinger, 2019; Grimminger y Klinger, 2021; Herrero-Izquierdo et al., 2022).

Así, se reconoce que la desafección política podría ser potencialmente detectada en la red social X, toda vez que el texto construye una realidad simbólica apoyada en las palabras y codificadas en un contexto socio político desde el cual se produce el sentido social de los contenidos, con representaciones de los actores e instituciones y dinámicas relacionadas con la democracia misma. Las acciones e interacciones expresan por su parte una compleja cultura de uso, cada vez más combinada con las lógicas algorítmicas predominantes.

Por su lado, los hashtags, palabras encerradas con numeral (#) en los contenidos publicados en X, son reconocidos como unidades temáticas de potencial análisis discursivo; su estudio, desde las posibilidades sociotécnicas para la descarga y organización de los mensajes, ya lleva una década de amplia revisión en distintos temas (Briceño-Romero y Bravo, 2022). Pano Alamán (2020) reconoce entre las funciones comunicativas del hashtag, las informativas, persuasivas, argumentativas o expresivas, según la interacción promovida entre los usuarios de las conversaciones derivadas de estas palabras en Twitter, acompañadas de un numeral (#), con códigos específicos compartidos por el enunciador y el enunciatario. La revisión de #Elecciones2022 desde una mirada comparada en distintos contextos electorales contribuiría con una nueva mirada a estos antecedentes.

3. Marco contextual

El año 2022 experimentó procesos electorales en cinco países de habla española, en diferentes fechas, desde febrero hasta octubre: Colombia, Costa Rica, España, México y Perú.

En América Latina, el calendario establecía un cronograma planificado que se llevó a cabo como estaba previsto en el año 2022 en Costa Rica y Colombia con votaciones para elegir presidente. En México se realizaron elecciones regionales, cuyo calendario coincidió para seis estados, además de otros procesos de elección municipal, generados

durante el primer semestre. Y, finalmente, Perú se sometió a un proceso de elecciones regionales y municipales en el último trimestre del año, en un ambiente que estuvo teñido por la intensificación de una nueva crisis institucional. Por su parte en Europa, España celebró en el año 2022 dos procesos electorales centrados en votar por los integrantes parlamentarios de las comunidades autónomas Castilla y León y Andalucía, en medio de una tensión discursiva entre partidos de izquierda y de derecha.

Estos países expresan procesos diversos de madurez democrática, con una trayectoria electoral fragmentada históricamente en el siglo XX, aunque fortalecida en el siglo XXI. Todos ellos han contado con un sistema de partidos tradicionales, aunque hacia los últimos años han experimentado nuevos movimientos políticos que se han insertado en la oferta electoral. Aunque con contextos particulares, estos escenarios electorales en países de habla hispana en los que la población tomó decisiones para elegir representantes locales, regionales y presidenciales, estimularon un intercambio discursivo centrado en el voto como indicador democrático, con la participación de actores e instituciones involucradas. La tabla 2 presenta un balance de las elecciones realizadas en el período y contexto estudiado.

Tabla 2. *Contexto electoral en los países de habla española en el año 2022*

Fecha	País	Proceso electoral	Contienda electoral
6 febrero	Costa Rica	Presidenciales 1era vuelta	Oficialismo-oposición Partido oficialista: Acción Ciudadana (Jorge María Figueroa) Oposición: Partido Progreso Social Democrático Rodrigo Chaves Robles
3 abril		Presidenciales 2da vuelta	
13 febrero	España	Cortes de Castilla y León	Partido de turno (centro izquierda) Partido Socialista Obrero Español (PSOE) Oposición (derecha): Partido Popular (PP) Izquierda radical neutralizada
19 junio		Parlamento Andalucía	
13 marzo	Colombia	Elecciones legislativas	Candidaturas antistablishment Movimientos políticos y alianzas Gustavo Petro Pacto Histórico: izquierda radical Rodolfo Hernández Candidatura Independiente / Liga Gobernantes Anticorrupción: derecha imprecisa
29 mayo		Presidenciales 1era vuelta	
19 junio		Presidenciales 2da vuelta	

(continúa)

Tabla 2. *Contexto electoral en los países de habla española en el año 2022 (continuación)*

Fecha	País	Proceso electoral	Contienda electoral
5 junio	México	Elecciones regionales (gobernadores, diputados y ayuntamientos)	oficialismo-oposición Partido oficialista de izquierda: MORENA Oposición partidos tradicionales: Partido Revolucionario Institucional (PRI)/centro derecha Partido de Acción Nacional/conservador
2 octubre	Perú	Elecciones regionales y municipales	Crisis institucional Crisis de partidos Movimientos regionales

Fuente: *creación propia.*

4. Metodología

La unidad de análisis de esta investigación ha sido el corpus derivado de mensajes en X encerrados en la etiqueta #Elecciones2022, cuya descarga fue generada desde el API REST y API STREAMING, por medio de librerías de Python con un script específico para los contenidos originales (no RT) y en español. Los datos fueron descargados de la red social X desde una cuenta de desarrollador adjudicada por esta empresa para fines académicos. El histórico de las descargas fue del 01 febrero al 20 de junio del año 2022, como un período exploratorio en un diseño transeccional.

El corpus fue sometido al siguiente tratamiento:

1. Caracterización de la etiqueta #Elecciones2022, según:
   a. Número de tuits por fecha / país
   b. Número de tuits por país (según geolocalización)
2. Clasificación automatizada de mensajes con la herramienta Perspective, tomando como referencia la categoría de:
   a. Toxicidad: 0.3 a 0.6.
   b. Toxicidad severa: 0,7 a 0,9 como máxima puntuación.
3. Clasificación manual de mensajes con toxicidad severa, por país:
   a. Tipo de desafección, según el centro del mensaje.

Tomando en cuenta la clasificación de cultura política propuesta en Briceño-Romero, Calderón-Benavides y Jurado (2022) los mensajes ubicados en la categoría de toxicidad severa, en una puntuación de 0.7 a 0.9, se sometieron a una clasificación manual de contenidos, desde la noción de desafección política, según el centro donde recae la toxicidad, como se explica en la tabla 3.

Tabla 3. *Categorías de análisis para la clasificación de contenidos centrados en la desafección política*

Tipo de desafección	Indicador	Descripción	Ejemplo Tuits*
Desafección política	Actores políticos	Se destaca la desafección centrada en personalidades políticas (candidatos, líderes de partidos o movimientos, o funcionarios públicos, generalmente con cargos relevantes, como presidente, gobiernos, alcaldes, concejales, diputados).	Aquí es donde uno ve de donde viene lo de dejar la mujer en la cocina o cuidando los hijos y el lenguaje soez hacia las trabajadoras sexuales. Viejo conchudo hpta.#Elecciones2022 #PetroPresidente2022@ petrogustavo @ingenierorodol @mario_delgado confirmando lo que ya sabíamos; un absoluto pendejø arrastrado que ni a su padrote pudo complacer. Bien hecho, batracio!!! Ni con el aparato gubernamental e ilícitos pudiste. #loser #Elecciones2022 @PartidoMorenaMx Figueres... populista de mierda!#Elecciones2022#EleccionesCR
Desafección institucional	Instituciones públicas Fuerzas Armadas Medios de comunicación Partidos Políticos Iglesia Otros	El discurso se centra en cualquiera de estas instituciones que funcionan en el sistema democrático: partidos políticos, poder legislativo (Congreso, Senado, concejos), poder ejecutivo (Presidencia), poder judicial (tribunales, Procuraduría, Fiscalía). Alcaldías, gobernaciones, Policía, fuerzas armadas, estructuras públicas del Estado (universidades, hospitales, etc.) y medios de comunicación.	@Registraduria En Colombia co no es un secreto que Uds otra vez se quieren robar las elecciones, por favor no lo hagan; no sean tan HIJUEPUTAS @petrogustavo #EleccionesColombia #Elecciones2022 #Registraduria

(continúa)

Tabla 3. *Categorías de análisis para la clasificación de contenidos centrados en la desafección política (continuación)*

Tipo de desafección	Indicador	Descripción	Ejemplo Tuits*
Cultura país	Identidad país Grupos poblacionales	Se destacan aspectos inherentes a rasgos de identidad del país o región. El texto asoma (fortalezas o debilidades) asociadas a los códigos que estructuralmente unen a un colectivo.	#Elecciones2022Independicen ese hijueputa cagadero ya. Más de 20 años arruinando este país. Paisas setenta hijueputas. Remalparidos pais de mierda no aprende me tienen harta hdp #Elecciones2022 Me cago en todos los que votaron por Rodrigo Chaves carepichas.#eleccionesCR2022
Ejercicio democrático	Democracia Elecciones	Desafección explícita hacia la democracia o a los valores que la identifican: elecciones, igualdad, separación de poderes, pluralidad de partidos, libertad de expresión, inclusión, participación.	Tuits ficticios**: No voy a votar porque no creo en ninguno de esos hjdp ladrones Prefiero una dictadura que esta democracia de mierX&D
Idea política		Se destaca como centro una desafección a las ideologías de izquierda o de derecha.	Andaluces de bien, mañana no olvidéis de meter las papeletas del PP y VOX en el mismo sobre, y que se jodan los comunistas.La unión hace la fuerza!#EleccionesAndalucia2022 #Elecciones2022 #Andalucía

*Los ejemplos son sacados del corpus analizado, presentados tal como fueron publicados.

**No se encontraron ejemplos reales en el corpus analizado, en esta categoría.

Fuente: *Propuesta adaptada de Briceño-Romero et al. (2022).*

5. Resultados

El universo de datos con mensajes de la etiqueta en Twitter #Elecciones2022, para un período de descarga del 01 febrero al 20 de junio del año 2022 fue de ciento quince mil cuatrocientos noventa y tres tuits (115.493).

5.1. Caracterización de la etiqueta

Tomando en cuenta la fecha de publicación de los mensajes, la tendencia de publicación de contenidos en Twitter en la etiqueta #Elecciones2022 destaca actividad continua durante el período analizado, aunque se reportan unos picos en fechas que corresponden a procesos clave de las elecciones, principalmente en: Colombia, España y México.

- La fecha en la que se reporta más actividad fue el 29 de mayo, con más de 35 mil tuits publicados. Esta fecha corresponde a la segunda vuelta electoral de las elecciones presidenciales en Colombia.
- En segundo lugar se reporta un pico de actividad el 13 de marzo, con más de 20 mil tuits. En esta fecha se realizaron las elecciones legislativas en Colombia.
- El 19 de junio se reporta un tercer pico de actividad, con cerca de 20 mil tuits publicados. En esta fecha se realizaron elecciones parlamentarias en la comunidad autónoma de Andalucía, España.
- Por último, se destaca un cuarto pico de actividad el 5 de junio, con más de cinco mil tuits. Esta fecha corresponde a las elecciones regionales en México.

En el gráfico 1 puede verse la tendencia de actividad en la etiqueta #Elecciones2022, con picos que evidencian su uso intenso en momentos electorales durante el período analizado.

Gráfico 1. *Tendencia actividad de #Elecciones2022 en el período analizado*

Fuente: *creación propia.*

Para identificar la actividad de la etiqueta por países, se revisó la información de geolocalización de las cuentas, con lo que se reporta lo siguiente: del corpus total, un grupo de 77 mil 169 tuits se encuentran geolocalizados, muestra que terminó siendo tomada para realizar la clasificación posterior. La distribución de los contenidos que fueron publicados explícitamente desde cuentas geolocalizadas reportan: Colombia fue el país con usuarios más activos en la etiqueta analizada, con 53.943 tuits, con una diferencia sustancial hacia el resto de los países. México ocupó el segundo lugar en la etiqueta, con 16.352 tuits. Mucho más alejados se encuentran: Costa Rica en tercer lugar, con 3.188 tuits. Y finalmente Perú en cuarto lugar, con 1.867 tuits y España con 1.819 tuits. Estos datos son resumidos en la tabla 4.

Los mensajes que han quedado dentro de la muestra final para la detección de desafección política en mensajes clasificados como altamente tóxicos han sido aquellos cuya información de geolocalización es explícita y atiende al país analizado, para garantizar que la discusión se encontrara dentro del contexto revisado. En este sentido, ha sido descartado el grupo de mensajes que no contaba con información de la geolocalización de los usuarios, corpus que merece una revisión aparte y complementaria a esta investigación. Se reporta esta condición como una limitación de la muestra analizada.

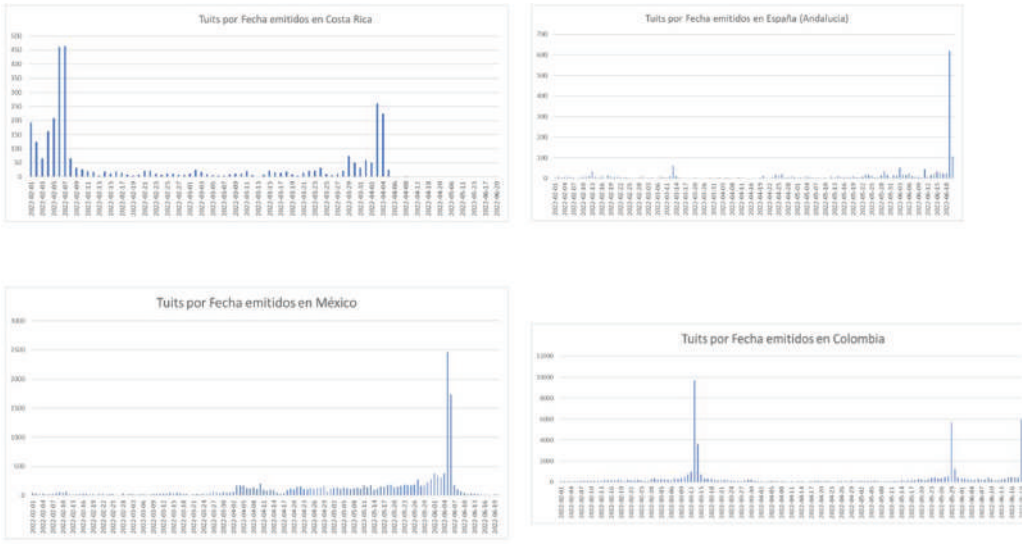
Los usuarios de X en cada país activaron la publicación de los contenidos de #Elecciones2022 en las fechas clave de cada proceso electoral, como puede verse en la comparación del gráfico 2.

Tabla 4. *Participación país en la etiqueta #Elecciones2022 en Twitter*

País	Número de tuits
Colombia	53.943
España	1.819
Perú	1.867
Costa Rica	3.188
México	16.352
Mensajes geolocalizados	77.169
Mensajes sin geolocalización	38.324
Corpus total	115.493

Fuente: *creación propia.*

Gráfico 2. *Comparación actividad de la etiqueta #Elecciones2022, por países*



Fuente: *creación propia.*

5.2. Mensajes tóxicos

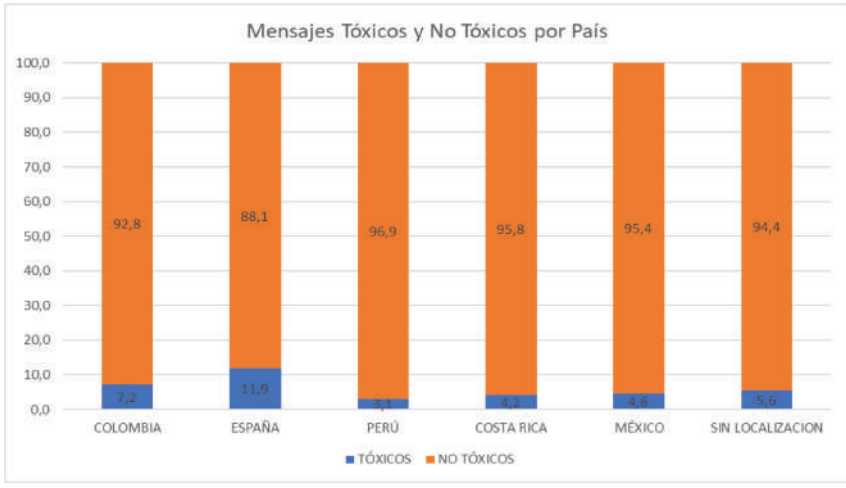
Tomando en cuenta la clasificación generada por Perspective, se reporta en la etiqueta #Elecciones2022 un porcentaje muy bajo de mensajes considerados tóxicos, por país, en una escala entre 0.3 y 0.9, un primer hallazgo que sugiere las etiquetas electorales más concentradas en procesos relacionados con la contienda y el acto electoral, con menos tendencia a la agresividad discursiva. No obstante, la muestra de mensajes con toxicidad severa resultó propicia discursivamente para la detección de la desafección política.

Como puede verse en el gráfico 3, proporcionalmente, el país que reportó mayor porcentaje de mensajes tóxicos fue España (11.9%); en segundo lugar, se reporta Colombia, con 7,2% de sus mensajes calificados como tóxicos. México y Costa Rica reportaron 4,6% y 4.2% de mensajes tóxicos, respectivamente. El país con menos porcentaje de contenidos calificados como tóxicos fue Perú, con 3.1%.

El gráfico 4 muestra la tendencia de toxicidad en los mensajes de la etiqueta #Elecciones2022, por países involucrados.

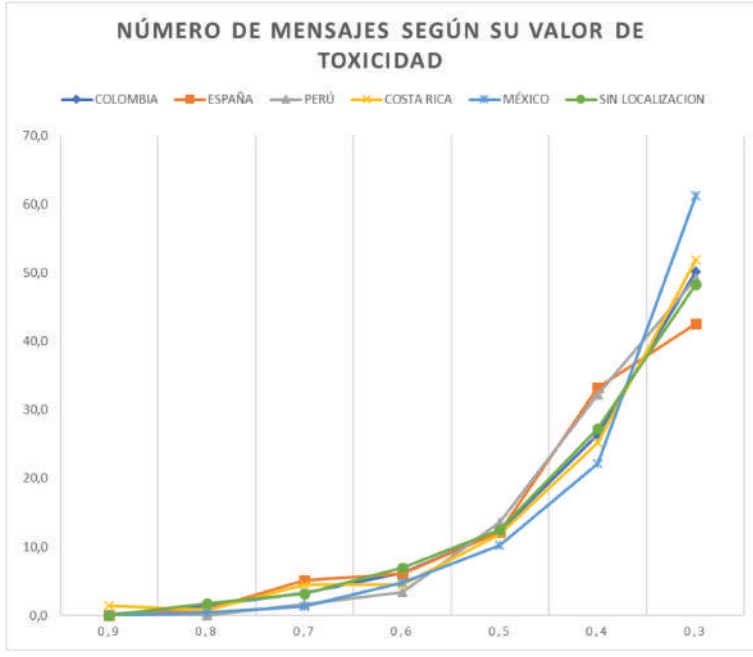
En general, los mensajes con toxicidad severa en los países analizados reportan baja viralidad y los usuarios mantienen perfiles poco relevantes, con un número

Gráfico 3. *Proporción de mensajes tóxicos, por país, en #Elecciones2022*



Fuente: *creación propia.*

Gráfico 4. *Nivel de toxicidad por países en #Eleccciones 2022*



Fuente: *creación propia.*

bajo de seguidores. Se destaca que no se reportaron perfiles de usuarios políticos entre quienes publicaron estos mensajes y en la revisión detallada de la bio, surgieron algunos perfiles sospechosos o con cuentas anuladas. Aunque esta condición merece una revisión en nuevas investigaciones, destacamos que, independientemente del perfil de usuarios, existe un contexto político que organiza potencialmente el sentido de la palabra alrededor del odio y la toxicidad, cuyo orden simbólico es aprovechado para capitalizar el descontento, bien sea voluntario o dirigido.

5.3. Desafección política

Para la detección de desafección política se realizó una clasificación humana de los mensajes tipificados con toxicidad severa, es decir aquellos con puntuación del 0.7 al 0.9 por la herramienta Perspective. Esta muestra agrupó un número de doscientos sesenta y tres (263) mensajes distribuidos entre cuatro países: Colombia (221 mensajes), Costa Rica (11 mensajes), España (16 mensajes) y México (15 mensajes). Perú no alcanzó a tener mensajes con toxicidad severa para su clasificación.

Se reporta que la muestra con mensajes clasificados como de toxicidad severa, es decir, con un puntaje adjudicado por la herramienta Perspective entre 0.7 y 0.9, expresa construcciones discursivas claras en cuanto a: agresividad del lenguaje, uso de mayúsculas, uso de emoticones y signos repetidos como los de admiración; además, estos contenidos suelen tener un centro claro sobre el cual recae la toxicidad, lo que permitió clasificar con facilidad el tipo de desafección política hacia actores políticos y hacia las instituciones. El uso de funciones conversacionales o llamados con el @usuario y el nombre explícito de las personas e instituciones aseguró la revisión de perfiles hacia los cuales se quería atacar o sobre las cuentas sobre las cuales recayó el lenguaje tóxico.

Ejemplos de mensajes con detección de desafección política.

> **Hacia actores políticos:**
> **Candidatos**
> Colombia
> Rodolfo Hernández no es sino un desgraciado hijueputa #SegundaVuelta #Elecciones2022 #EleccionesColombiaEjemplos

El cacas de la guerrila siendo el @petrogustavo de las elecciones. Jajajajaja esto para los #PetroMeQuiereMucho no les parece #CORRUPCION ese hp es lo más corrupto de Colombia y aparte socialista. Este es un mentiroso de primera.

España
#EleccionesAndaluzas@Macarena_Olona A mamarla!Ah no...que la masturbación y la educación sexual te la pelaPues a joderte y a vender el vestido de flamenca en walapop#Elecciones2022 #EleccionesAndaluzas #19JAndalucia #19JAndaluciaL6#Andalucia19J

Fracaso de la basura fascista de @vox_es y la putrefacta @Macarena_Olona pensabais que ibais a gobernar jajajsjsaja -336.817 que han despertado a vuestro ridiculo cerdos. JAMÁS #Elecciones2022 #EleccionesAndaluzas #voxbasura #Antifascista #amamarla #Vaxura #España

Costa Rica
Un agresor o un acosador sexual será el próximo presidente de Costa RicaQUE MIERDA DE GENTE #Elecciones2022 #EleccionesCR #costarica

Rodrigo Chaves siendo un pedófilo, machista, asqueroso y todo lo malo que puede haber es el nuevo presidente de Costa Rica. odio a cada uno de los que votó por él.#EleccionesCR #Elecciones2022 #acosador

México
Un imbecil defendiendo a otro imbecil así la 4ta#LopezBasuraPresidencial #MorenaDestruyendoAMexico #Elecciones2022 #ElImbecilDelPalacio

Partidos y políticos tradicionales
Colombia
Lo que más me alegra de estas #Elecciones2022 es la quemada del Nuevo liberalismo. Maldito par de porquería, sátrapas setenta hptas delfines vividores del estado @juanmanuelgalan @CarlosFGalan Están ahogaoooo perros

Boletín 7!!! 47.18% GONORREEEAAAA EN TU CARA URIBE DE MIERRRRDA AAAAAAAAAAA!!!

México
El #PRI siendo el #pri con 153 periodistas asesinado en México, #Alito #AlitoMoreno con estas declaraciones, #hijodeputa#periodistas #Mexico #partidospolíticos #Mexico

En el caso de la detección de la desafección plasmada como una frustración hacia la cultura país, los mensajes estuvieron más concentrados en la fuerza de la palabra y en general hacían alusión al nombre del país, región o grupo poblacional sobre el cual recaía la toxicidad. En el caso de toxicidad expresada hacia ideologías, las construcciones fueron también explícitas. Ejemplos de desafección:

Cultura país:
Colombia
Los odio país de mierda.

ME DUELES COLOMBIAAAA💔 💔 qué país tan hijueputa!!

Definitivamente prima la ignorancia, no leer propuestas, no saber del pasado del candidato, no saber ni mierda, gente de mierda 🤦🏻‍♀️ Qué angustia #EleccionesColombia #Elecciones2022

Hacia grupos poblacionales con afición política:
Colombia
PEROO QUIENES SON LOS IMB3CIL3S Q VOTAN POR EL VIEJO ESEEEE #Elecciones2022

Y así se podrán reproducir por todo el país!!! porque la cantidad de hijueputas que hay apoyando al malandro de @FicoGutierrez se encuentran en cloacas…

España
Lo que pasa en Andabasuria (andalucia) es el fiel reflejo de la puta mierda que es Expaña.#Elecciones2022 #EleccionesAndaluzas #Elecciones19J

Ideología
España
Socialistas de mierda

En general, la clasificación de desafección política de la muestra reportó un número predominante de mensajes con toxicidad centrada en actores políticos (109 mensajes; 41.4%), así como también en la cultura país (108 mensajes; 41,1%). En tercer lugar, muy alejado de la tendencia, se encuentra un grupo de mensajes que expresan desafección institucional (26 mensajes; 9.9%). Sólo dos (2) mensajes (0.8%) fueron clasificados como desafección hacia ideas políticas. Se reporta un

total de dieciocho (18) mensajes (9.9%) cuya redacción resultó ambigua y, por tanto, de difícil clasificación.

La tabla 5 detalla la clasificación de desafección por país en esta muestra. La desafección detectada en los mensajes con toxicidad severa de la muestra se expresa desde una construcción discursiva casi explícita. Sólo en pocos mensajes se generó ambigüedad o confusión relacionada con la falta de contexto.

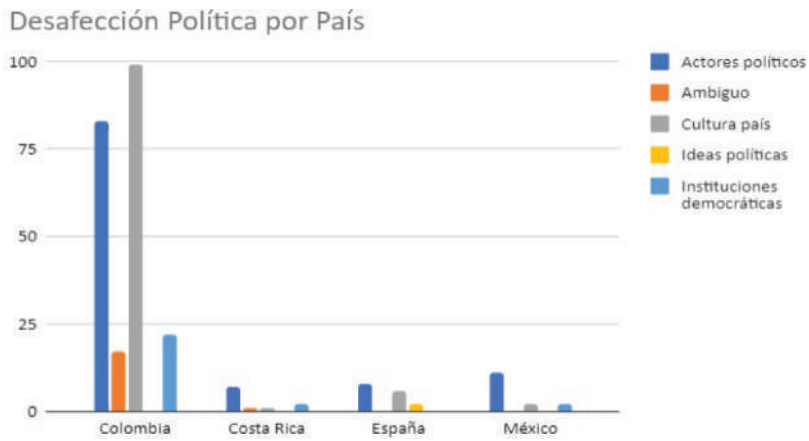
Tabla 5. *Desafección política detectada en los mensajes tóxicos en #Elecciones2022*

CLASIFICACIÓN DE LA DESAFECCIÓN POLÍTICA E INSTITUCIONAL EN #ELECCIONES2022					
País	Actores políticos	Cultura país	Instituciones	Ideas políticas	Ambiguo
Colombia	83	99	22		17
Costa Rica	7	1	2		1
España	8	6		2	
México	11	2	2		2

Fuente: *creación propia.*

En el gráfico 5 puede verse la tendencia de desafección política, según el centro hacia el cual recae el malestar, en los mensajes estudiados.

Gráfico 5. *Desafección política por país detectada en mensajes con alta toxicidad en #Elecciones2022*



Fuente: *creación propia.*

6. Consideraciones de cierre

Esta experiencia exploratoria refuerza que el hashtag #Elecciones2022 se inscribe dentro de la función contextualizadora, con intención informativa (Pano Alamán, 2020) en Twitter (X), dado que su actividad emerge con gran fuerza hacia momentos clave electorales de cada país en donde la comunidad de usuarios publica sus contenidos relacionados con el voto y los posibles resultados electorales.

En este sentido, la desafección política encontrada surge de emociones puntuales negativas en las que se expresa rechazo a la dicotomía cambio-continuidad en el poder político, desde el proceso electoral. En este caso los resultados asomaron una toxicidad discursiva centrada en: frustración ante el cambio político (Colombia), la dualidad oficialismo–oposición (México y Costa Rica), y fuerzas ideológicas derecha–izquierda (España), aunque muestras más grandes en exploraciones futuras podrían confirmar tendencias al respecto.

La revisión humana de los mensajes clasificados con toxicidad severa por la herramienta Perspective (0.7 a 0.9) sugiere que esta tipificación discursiva es eficiente para la detección de la desafección política, entendiendo que existen evidencias lingüísticas de sentimientos negativos explícitos y, en su mayoría, se reconoce el centro sobre el cual recae el malestar, expresado en el discurso textual.

No obstante, comprendiendo que la tendencia de toxicidad en la etiqueta analizada ha sido baja, se sugieren análisis futuros en etiquetas activistas o relacionadas con las campañas de movimientos que defiendan explícitamente la representación de partidos o candidatos específicos, con el fin de evaluar estrategias discursivas que podrían promover el descontento como capital electoral. Así, aunque este caso de estudio no arrojó porcentajes sustanciales de mensajes considerados con toxicidad severa por la herramienta Perspective, la aplicación de la metodología en investigaciones futuras en hashtags con otras funciones comunicativas (argumentativas, persuasivas), podría ayudar al acopio de expresiones discursivas que deriven en entrenamientos, con el fin de complementar la detección de la desafección política desde aprendizajes automatizados.

En términos específicos, los resultados demuestran la fuerza explícita de la palabra escrita tipificada con alta toxicidad, para detectar desafección principalmente hacia los actores políticos, en gran parte candidatos de las contiendas electorales o figuras

públicas de partidos, ideologías o movimientos a los cuales se les ataca. Igualmente se logra detectar construcciones estereotipadas de grupos poblacionales a los cuales se les señala por su cultura política. Así, en el contexto analizado la política parece impulsar grupos sensibles de ser señalados y atacados por sus aficiones históricas. El ataque en este sentido acude a debilitar la imagen de estos actores, con descalificativos que encierran un sentido social en códigos acumulados simbólicamente.

Alineados con la noción de sociolingüística ciudadana (Bridges, 2021), el lenguaje digital que emerge en X podría ayudar a configurar a mediano y largo plazo un panorama de creencias sobre el sistema político en determinados contextos y contribuir así a un seguimiento de cultura país desde una perspectiva computacional, útil para comprender la tendencia de rechazo a los sistemas políticos. Continuar estudios como estos, desde una perspectiva longitudinal podría preparar muestras potencialmente entrenables para la detección automatizada de la desafección política, como complemento de otras metodologías tradicionales.

Contribución de autoras

Ysabel Briceño-Romero: Conceptualización, Conservación de los datos, Análisis formal, Investigación, Metodología, Supervisión, Validación, Visualización, Redacción – borrador original, Redacción – revisión y edición. **Liliana Calderón-Benavides:** Metodología, Validación, Visualización, Redacción – revisión y edición. Los autores han leído y están de acuerdo con la versión publicada del manuscrito.

Conflictos de interés

Las autoras declaran no tener conflicto de interés.

Financiación

Este artículo se deriva del proyecto de investigación Unab titulado “Observación electoral de la opinión pública en 280 caracteres” (código 26764718).

Referencias

Almond, Gabriel; & Verba, Sidney. (1963). *The civic culture*. Princeton University Press.

Amores, Javier; Blanco-Herrero, David; Sánchez-Holgado, Patricia; & Frías-Vázquez, Maximiliano. (2021). Detectando el odio ideológico en Twitter. Desarrollo y evaluación de un detector de discurso de odio por ideología política en tuits en español. *Cuadernos.Info, (49),* 98–124. https://doi.org/10.7764/cdi.49.27817

Arcila-Calderón, Carlos; Blanco-Herrero, David; & Valdez-Apolo, María Belén. (2020). Rechazo y discurso de odio en Twitter: Análisis de contenido de los tuits sobre migrantes y refugiados en español. *Revista Española de Investigaciones Sociológicas, 172*, 21-40. https://doi.org/10.5477/cis/reis.172.21

Arcila-Calderón, Carlos; Sánchez-Holgado, Patricia; Quintana-Moreno, Cristina; Amores, Javier; & Blanco-Herrero, David. (2022). Hate speech and social acceptance of migrants in Europe: analysis of tweets with geolocation//Discurso de odio y aceptación social hacia migrantes en Europa: análisis de tuits con geolocalización. *Comunicar*, *30*(71), 21-35.

Bollen, Johan; Mao, Huina; & Pepe, Alberto. (2021). Modeling Public Mood and Emotion: Twitter Sentiment and Socio-Economic Phenomena. *Proceedings of the International AAAI Conference on Web and Social Media, 5(1),* 450-453. https://doi.org/10.1609/icwsm.v5i1.14171

Briceño-Romero, Ysabel, & Bravo Bautista, Luz. (2022). La movilización social en entornos digitales: una revisión de la producción científica en español en el siglo XXI. *Reflexión Política, 24(49),6-20*. [fecha de Consulta 24 de Junio de 2025]. ISSN: 0124-0781. https://www.redalyc.org/articulo.oa?id=11076003001

Briceño-Romero, Ysabel; Calderón-Benavides, Liliana; & Jurado, Miguel. (2022). Clasificación de sentimientos en Twitter desde la noción de cultura política: una

revisión discursiva en el escenario electoral de Colombia. *Anuario Electrónico De Estudios En Comunicación Social "Disertaciones", 18(1)* https://hdl.handle.net/20.500.11912/10444

Bridges, Judith. (2021). Explaining -splain in digital discourse. *Language Under Discussion. Vol. 6, Issue 1 (July 2021),* pp. 1–29.

Capella, Joseph; & Jamieson Hall, Kathleen (1997). *Spiral of Cynicism: The Press and the Public Good*. Nueva York: Oxford University Press.

Corporación Latinobarómetro. (2024). Informe 2023. La recesión democrática de América Latina. Ministerio de Asuntos Exteriores y Cooperación, Banco Interamericano de Desarrollo (BID) y AECID. España.

Cucurull, Irina; & Aragó Navarro, Bernart. (2023). Odio en Twitter: la intersección entre género y racismo. Informe NovAct. España.

Díez-Gutiérrez, Enrique; Verdeja, María; Sarrión-Andaluz, José; Buendía, Luis; & Macías-Tovar, Julián. (2022). Political hate speech of the far right on Twitter in Latin America. [Discurso político de odio de la ultraderecha desde Twitter en Iberoamérica]. *Comunicar, 72,* 101-113. https://doi.org/10.3916/C72-2022-08

Fontaneda, Javier; & Sánchez-Vítores, Irene. (2018). La desafección en las urnas: las elecciones generales de 2015 en España / Disaffection at the Ballot Box: The 2015 General Election in Spain. *Reis: Revista Española de Investigaciones Sociológicas, 161,* 41–62. https://www.jstor.org/stable/44841756

Grimminger, Lara; & Klinger, Román. (2021). Hate Towards the Political Opponent: A Twitter Corpus Study of the 2020 US Elections on the Basis of Offensive Speech and Stance Detection. In Proceedings of the Eleventh Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, pages 171–180, Online. Association for Computational Linguistics. https://aclanthology.org/2021.wassa-1.18/

Herrero Izquierdo, Jacobo; Reguero Sanz, Itziar; Berdón Prieto, Pablo; & Martín Jiménez, Virginia. (2022). La estrategia del odio: polarización y enfrentamiento partidista en Twitter durante las elecciones a la Asamblea de Madrid de 2021. *Revista Prisma Social, (39),* 183–212. https://revistaprismasocial.es/article/view/4829

Horta Ribeiro, Manoel; Pedro, Calais; Yuri, Santos; Virgilio, Almeida; & Wagner Meira, Jr. (2018). "Like Sheep Among Wolves": Characterizing Hateful Users on Twitter. In *MWSDM*.

Hosseini, Hossein; Kannan, Sreeram; Zhang, Baosen; & Poovendran, Radha. (2017). Deceiving Google's Perspective API Built for Detecting Toxic Comments. arXiv:1702.08138

Jain, Brown; Chen, Jeffery; Neaton, Erin; Baidas, Mohammad; Dong, Ziqian; Gu, Huanying; & Sertac, Nabi. (2018). Adversarial Text Generation for Google's Perspective API. *In CSCI*.

Keller, Tobías; & Klinger Ulrike. (2019). Social Bots in Election Campaigns: Theoretical, Empirical, and Methodological Implications. *Political Communication 36*, 1 pp. 171–189. https://doi.org/10.1080/10584609.2018.1526238

LAPOP. (2023). *Pulse of Democracy*. Lupu, Noam, Mariana Rodríguez, Carole J. Wilson, and Elizabeth J. Zechmeister (Eds.) Nashville, TN.

Lees, Alyssa; Tran, Vinh; Tay, Yi; Sorensen, Jeffrey; Gupta, Jai; Metzler, Donald; & Vasserman, Lucy. (2022). A New Generation of Perspective API: Efficient Multilingual Character-level Transformers. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '22)*, August 14–18, Washington, DC, USA. ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3534678.3539147

Marien, Sofie. (2011). Measuring Political Trust Across Time and Space. En Hooghe M., Zmerli S. (Eds.), Political Trust. Why Context Matters. (pp. 13-46). Colchester: ECPR Press https://ssrn.com/abstract=2539667

Megías Collado, Adrián. (2020). Una década de crisis desafecta: los cambios en su naturaleza. *Revista Española de Investigaciones Sociológicas, 169*, pp. 103-122. https://dx.doi.org/10.5477/cis/reis.169.103

Megías Collado, Adrián; & Moreno, Cristina. (2022). La desafección política en los países del entorno europeo español. *REIS. Revista Española de Investigaciones Sociológicas Núm. 179,* pp. 103-124 https://doi.org/10.46661/revintpensampolit.10941

Pano Alamán, Ana. (2020). La política del hashtag en Twitter. *En Vivat Academia, n.º 152 (septiembre),* pp. 49-68. https://doi.org/10.15178/va.2020.152.49-68

Perspective developers. (s/f). Attributes & Languages https://developers.perspectiveapi.com/s/about-the-api-attributes-and-languages?language=en_US

Piñeiro-Otero, Teresa; & Martínez-Rolán, Xabier (2021). Eso no me lo dices en la calle. Análisis del discurso del odio contra las mujeres en Twitter. *Profesional de la información, v. 30, n. 5,* e300402. https://doi.org/10.3145/epi.2021.sep.02

Pnud. (2022). *Gobernanza, Democracia y Desarrollo en América Latina y El Caribe*. https://www.undp.org/es/latin-america/publications/gobernanza-democracia-y-desarrollo-en-america-latina-y-el-caribe

Qayyum, Hina; Hao Zhao, Benjamin; Wood, Ian; Ikram, Muhammad; Kourtellis, Nicolas; & Ali, Mohamad. (2023). A longitudinal study of the top 1% toxic Twitter profiles. *In Proceedings of the 15th ACM Web Science Conference 2023 (WebSci '23)*. Association for Computing Machinery, New York, NY, USA, 292–303. https://doi.org/10.1145/3578503.3583619

Raghad, Alshalan; Hend, Al-Khalifa; Duaa, Alsaeed; Heyam, Al-Baity; & Shahad, Alshalan. (2020). Detection of Hate Speech in COVID-19–Related Tweets in the Arab Region: Deep Learning and Topic Modeling Approach. *Journal Medical Internet Research 22,* 12.

Torcal, Mariano; & Montero, José Ramón. (2006). Political Disaffection in Comparative Perspective. En Mariano Torcal y Jospe Ramón Montero (eds.), Political Disaffection in Contemporary Democracies: Social capital, Institutions, and Politics (pp. 3-19). Londres y Nueva York: Routlegde.

Valgarðsson, Viktor; & Devine, Daniel. (2022). What Satisfaction with Democracy? A Global Analysis of "Satisfaction with Democracy" Measures. *Political Research Quarterly, 75(3)*, 576-590. https://doi.org/10.1177/10659129211009605

Notas

[1] Pueden verse detalles de la herramienta acá: https://developers.perspectiveapi.com/s/about-the-api-attributes-and-languages?language=en_US